



Deep learning provides high accuracy in automated chondrocyte viability assessment in articular cartilage using nonlinear optical microscopy: supplement

XUN CHEN,^{1,2} YANG LI,¹  NICOLE WYMAN,¹ ZHENG ZHANG,¹ HONGMING FAN,¹ MICHAEL LE,¹ STEVEN GANNON,¹ CHELSEA ROSE,¹ ZHAO ZHANG,¹ JEREMY MERCURI,¹ HAI YAO,¹ BRUCE GAO,¹ SHANE WOOLF,³ THIERRY PÉCOT,⁴ AND TONG YE^{1,5,*} 

¹Department of Bioengineering, Clemson University, Clemson, SC 29634, USA

²Current address: Institute of Medical Photonics, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing 100083, China

³Department of Orthopedic, Medical University of South Carolina, Charleston, SC 29425, USA

⁴Hollings Cancer Center, Medical University of South Carolina, Charleston, SC 29425, USA

⁵Department of Regenerative Medicine and Cell Biology, Medical University of South Carolina, Charleston, SC 29425, USA

*ye7@clemson.edu

This supplement published with The Optical Society on 16 April 2021 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.14347040>

Parent Article DOI: <https://doi.org/10.1364/BOE.417478>

Deep Learning Provides High Accuracy in Automated Chondrocyte Viability Assessment in Articular Cartilage Using Nonlinear Optical Microscopy: supplemental document

S1. Schematics of the U-Net architecture for chondrocyte segmentation.

Here we introduce the image preparation, chondrocyte detection/segmentation, and post-processing steps using U-Net based convolutional neural networks as shown in Fig. S1. The details are described below.

Image preparation. Before segmentation, three major steps of preprocessing were performed on raw images. First, original SHG images in grayscale were inverted to set the intensity of cellular areas to high values. Second, the inverted images were filtered with opening/closing morphological operations for background removal and edge sharpening. This preprocess step improved the overall performance (Dice parameters) of cell segmentation. Lastly, images and related image masks were converted from the TIFF image format to the NPY (used in Python) numerical format before being input into the deep learning program. The input image set was randomly split into the training (3200 images) and test (800 images) sets with a ratio of 8:2. Among the 3200 training images, 640 images were used as validation images to confirm the best performance of the trained model.

U-Net chondrocyte detection and segmentation. The output of U-Net model was a series of binary masks, each of which contained only one cellular area, as shown in Figure S1(A). The cell detection algorithm was the weighted loss categorical cross-entropy as a loss function in deep learning and optimization. By using the predicted value of categorical cross-entropy, the U-Net model identified N number of cellular areas. The architecture of the U-Net model is shown in Figure S1(B), and typical loss curves for training and validation are shown in Figure S1(C). To confirm the highest IoU of segmentation, the outputs of the U-Net model were compared with the test/validation sets while varying the network. The U-Net network used the following parameters: batch size (16), kernel size (3*3 down convolution and 2*2 up convolution), specified activation filter (32, 64, 128, 256, and 512), specified loss function (binary crossentropy) and controlled epoch loop number (100).

Post-processing. Two steps are proceeded on the segmentation output. First, separated cell clusters of TPAF/SHG images were cropped using the segmented masks from the U-Net model. Cropped cell clusters were RGB images containing only cells for multivariate classification. Second, each segmented region of interest (ROI) was saved for final viability labeling with the outline of chondrocytes.

S2. Schematics of AlexNet architecture for chondrocyte multivariate classification

Here, we introduce the image preprocessing, chondrocyte multivariate classification, and post-processing steps using AlexNet based convolutional neural networks, as shown in Fig S2. The details are described below.

Preprocessing. The segmented clusters of TPAF/SHG images were pasted in black-grounded images with the size of 100×100 pixels. Then, images were resized to 256×256 pixels to build the training and testing sets. The training set was augmented with image rotation at 45°, 90°, 135°, 180°, 225°, 270°, and 315°. These images with related image categorical IDs were converted from image TIFF format to numerical NPY format NPY for the data input. This image data set contains 8280 cells with 5904 live cells among them. The input image set was randomly split into the training/validation sets (6624 cells in total with 4724 live cells) and test sets (1656 cells in total with 1180 live cells) with a ratio of 8:2. Furthermore, a validation set, containing 1324 cells in total with 944 live cells, was used to confirm the best performance of the trained model. Different segmentation algorithms have different segmentation results. Using UW, UN, WS, and ATB, the total number of sub-images segmented were 6256, 5072, 4384, and 3344, respectively. Fig. S3 summarizes the number of sub-images that contain different numbers of cells or different numbers of live cells in ground truths obtained by using different segmentation methods.

CNN chondrocyte multivariate classification. AlexNet convolutional neural network was used in the classification. As shown in Figure S2(B), this network included three convolution layers and two inner product layers. Following each convolution layer, a max-pooling and non-linear ReLU layer downsized the feature map gradually. A TPAF/SHG image with 256×256×3 pixels, similar to the one shown in Figure S2(A), was the input. The output layer output the probability of each category for an input image. Two independent CNNs were implemented for performing two 1D classifications with five categories (0, 1, 2, 3 and 4), one for the live cell number, and the other for the total cell number of input images in Fig. S2(A). The two independently trained AlexNet models with loss curves in Figure S2(C) were saved and loaded for categorical prediction and optimization. To confirm the highest accuracy of classification, the outputs of AlexNet models were compared with the test/validation sets while varying the networks. The AlexNet networks used the following parameters: batch size (64), kernel size (4*4), number of convolution kernels (32), specified activation filter (relu), specified loss function (categorical crossentropy) and controlled epoch loop number (100).

Post-processing. There are two ways to show the results of viability assessment: viability ratio and automated labelling of cartilage images. First, live/total cell classifications determined the viability ratio of the cartilage TPAF/SHG images for viability assessment. Secondly, live/total cell classifications of each segmented cluster were saved for final viability labelling with pseudo colors (red circles labeled dead cells, green circles labeled live cells, yellow circles labeled clusters of live and dead cells, magenta circles labeled clusters of non-cells), as shown in Fig. 7.

S3. Imaging microscope and sample holder

Immediately after euthanasia, the whole hind leg tibias male Sprague-Dawley rats were excised. The cartilage surfaces were exposed while kept intact on the long bone. The bone was held stable with a customized clamp (Fig. S4(C)) while resting in a glass-bottomed petri dish and immersed in Dulbecco's Phosphate-buffered Saline (DPBS, Corning Inc.) for microscopy imaging. The cartilage tissues were imaged with a commercial multiphoton optical microscope (Fig. 4(A); Olympus FV1200) equipped with a femtosecond pulsed laser (Mai Tai DeepSee, Newport Corp.) for two-photon excited fluorescence and SHG imaging. Two GaAsP photomultiplier tubes (PMTs) were equipped for optical signal detection in two channels simultaneously. For two-photon excited fluorescence microscopy, the wavelength of

excitation laser was tuned at 740 nm, which excites the endogenous fluorescent coenzymes of NAD(P)H and FPs at the same time. The fluorescence was collected with a 30 \times , silicone oil immersion Objective Lens (UPLSAPO30XS, Olympus UIS2) with a NA of 1.05 and then separated at 570 nm with a dichroic mirror (DM). Subsequently, the fluorescence passed through a violet (420—460 nm) and red (575—630 nm) bandpass filter and detected in two channels respectively. The laser wavelength was tuned to 860 nm for SHG imaging and detected with the violet channel. The frame size of acquired images were kept at 1024 \times 1024 pixels, corresponding to a field of view of 423 μ m \times 423 μ m. The pixel dwell time was set at 2- μ s with a Kalman filter integration of 2 lines. The z-stack images were collected at 1 μ m/slide for 50-slides starting from the surface downward into the tissue.

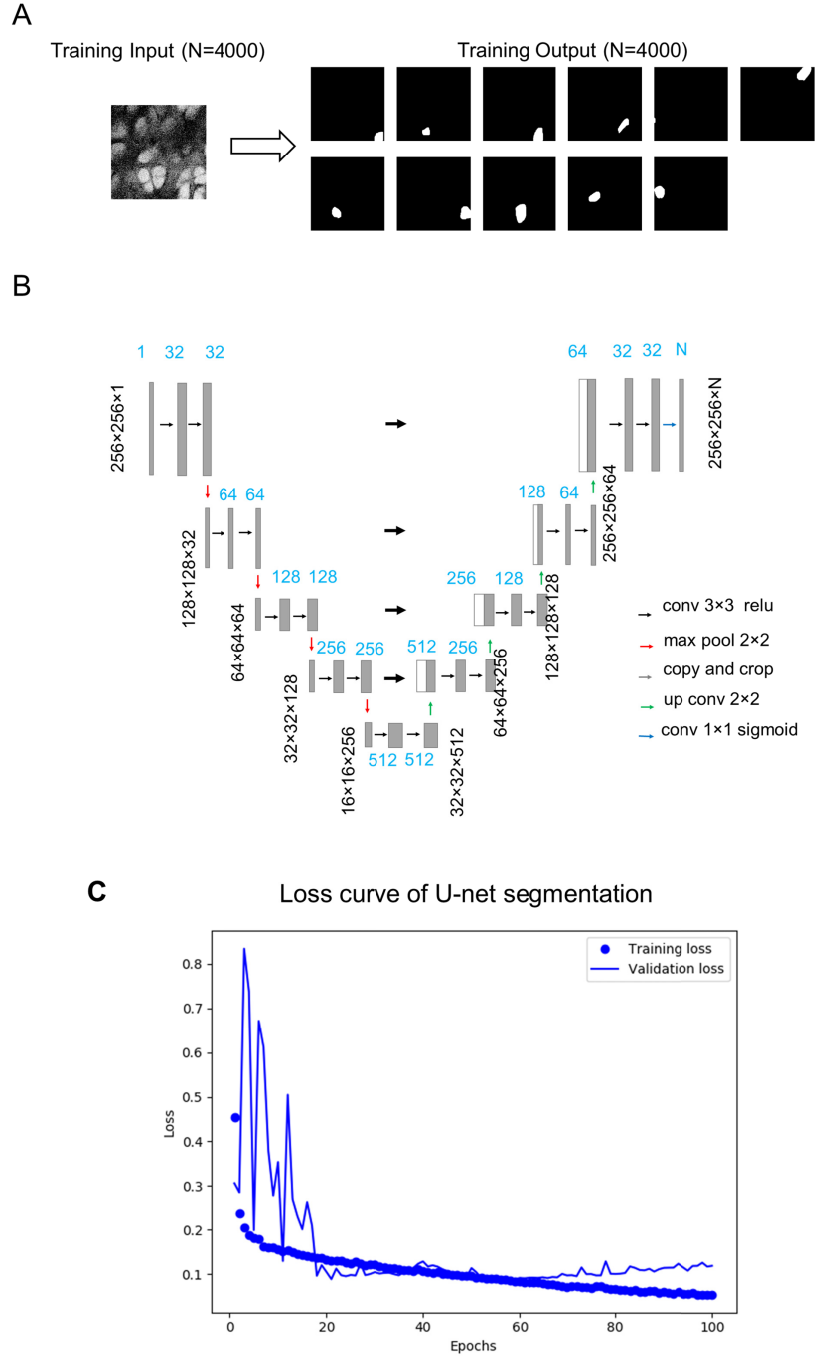


Fig. S1. U-Net architecture for chondrocyte segmentation. (A) Typical training input (N=4000) of cartilage SHG images and training output (N=4000) of artificial masks of U-Net model. (B) Neural networks of U-Net model using pixelwise loss weights segmentation. (C) Loss curves of training and validation for the U-Net segmentation model.

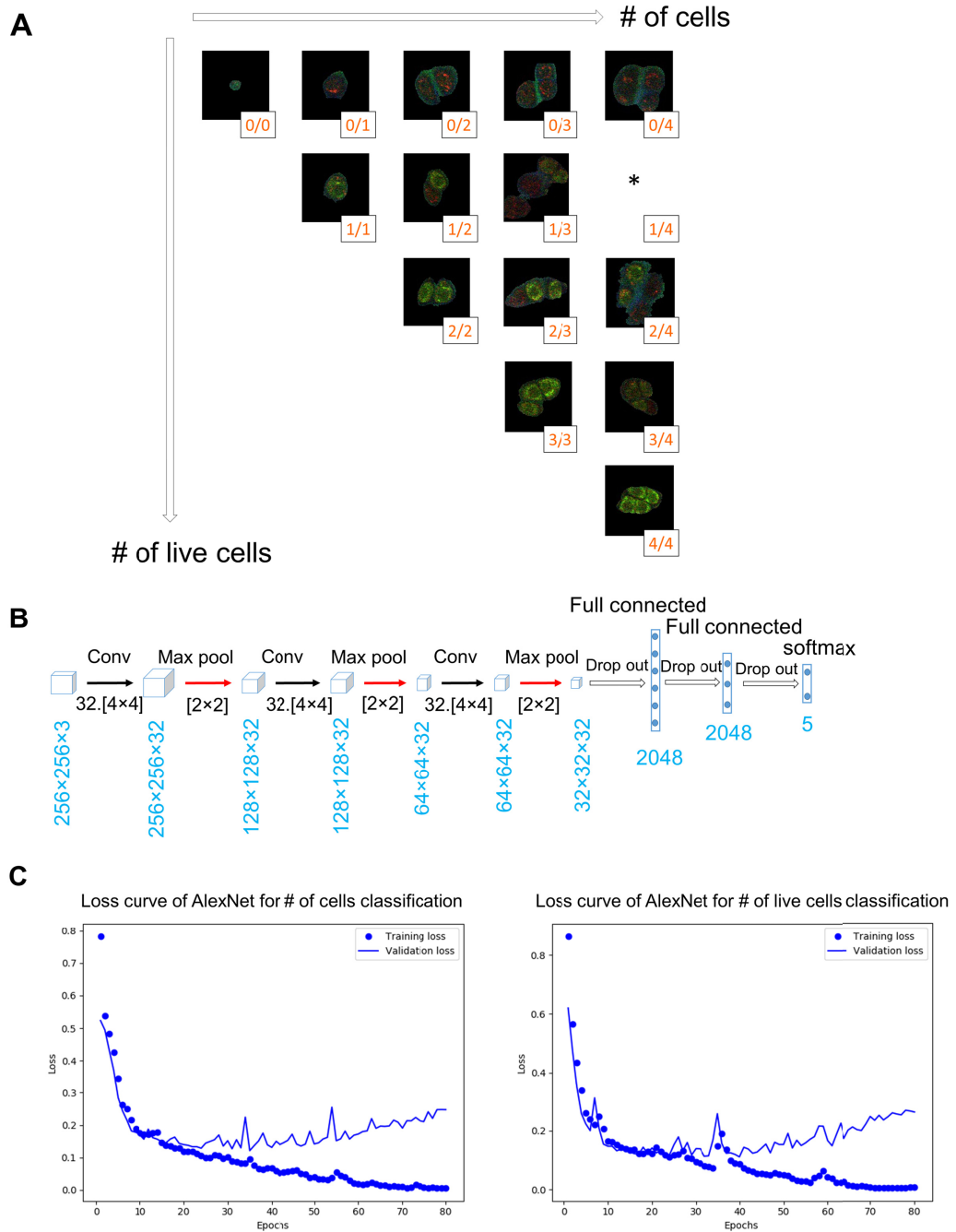


Fig. S2. Schematic of AlexNet architecture for chondrocyte multivariate classification. (A) Typical training input of cartilage SHG/TPAF images and training output of artificial counts of two separate AlexNet models. (B) Neural networks of AlexNet model using multivariate classification. (C) Loss curves of training and validation of two separate AlexNet classification models: one for total number of cells and the other for number of live cells classifications.

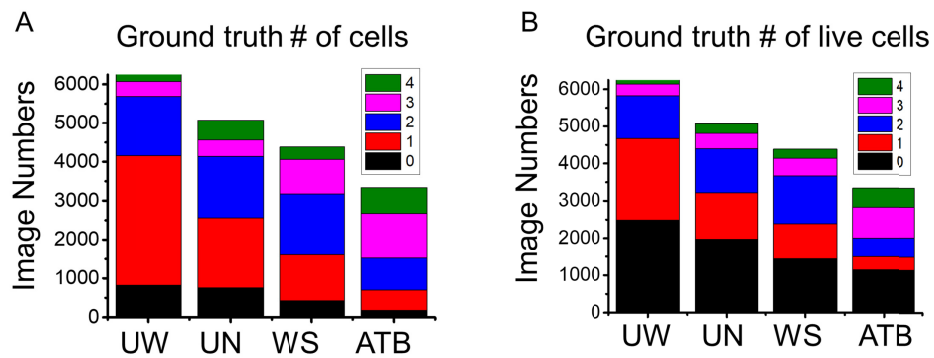


Fig. S3. Portions of sub-images containing different number (0-4) of live cells or cells in total in the dataset used as ground truth for the chondrocyte classification. No cell numbers more than 4 were found. (A) The total cell number obtained from each segmented image using different segmentation methods. (B) The live cell number obtained from each segmented image using different segmentation methods.

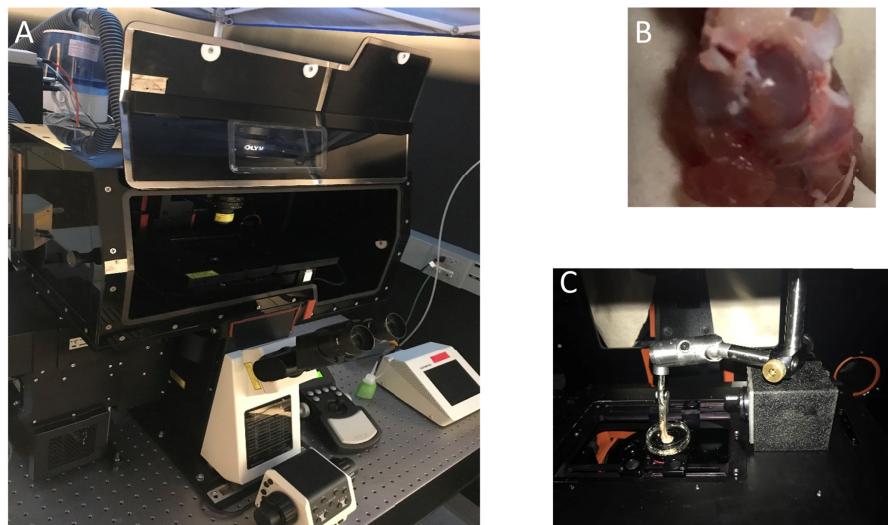


Fig. S4. (A) The multiphoton laser scanning microscope used in this study. (B) A photo of the exposed cartilage intact on a tibia condyle. (C) A photo of the customized clamp for holding tibias during imaging.